

FORSCHUNGSZENTRUM JÜLICH GmbH
Zentralinstitut für Angewandte Mathematik
D-52425 Jülich, Tel. (02461) 61-6402

Interner Bericht

UniTree® unter AIX

Testbericht

Lothar Wollschläger

KFA-ZAM-IB-9313

Oktober 1993
(Stand 4. Oktober 1993)

Zusammenfassung

UniTree wurde im Juli 1993 in der Version für AIX auf einer RS/6000 zusammen mit einem EXABYTE Roboter (EXB-120 CHS) auf Funktion, Leistungsfähigkeit und Zuverlässigkeit untersucht. Bei der Software handelt es sich um die AIX-Portierung von UniTree durch die Firma ACSC, die in Deutschland von M+S Elektronik vertrieben wird.

Die Untersuchungen haben ergeben, daß die Software ohne Fehler arbeitet und die zugesagten Funktionen erfüllt. UniTree fehlen jedoch wesentliche Administrator- und Recovery-Funktionen. UniTree ist nicht robust gegenüber Bandfehlern. Da der Zugriff zu UniTree-Daten über NFS oder FTP erfolgt, schränken alle Schwächen dieser Komponenten die Nutzung von UniTree als allgemeinen Fileserver ein.

Inhaltsverzeichnis

Zusammenfassung	iii
Verzeichnis der Abbildungen	vii
1 Einführung	1
1.1 UniTree	1
1.2 EXB-120 Cartridge Handling Subsystem	1
2 UniTree Installation	3
2.1 Aufwand	3
2.2 AIX Abhängigkeiten	3
2.3 Benutzereinrichtung	3
3 UniTree Benutzerschnittstelle	5
3.1 Allgemeines	5
3.2 Lokaler Zugang	5
3.3 FTP Zugang	5
3.4 NFS Zugang	5
3.5 Vergleich FTP — NFS	6
4 UniTree Systemadministration	9
4.1 Konfiguration	9
4.2 Benutzerverwaltung	9
4.3 Verwaltung der Kopien	9
4.4 Bandkompaktierung	9
4.5 Reinigung der Laufwerke	10
4.6 Sicherung der UniTree Datenbanken	10
4.7 Beenden von UniTree	10
5 UniTree Stabilität	13
6 UniTree Error Recovery	15
6.1 Bandfehler	15
6.2 Systemzusammenbruch	15
7 UniTree Performance	17
7.1 interne Performance	17
7.2 externe Performance	18
8 EXB-120 Cartridge Handling Subsystem	21
8.1 EXB-120 CHS Stabilität	21
8.2 EXB-120 CHS Performance	21
8.3 EXB-120 CHS Roboterbefehle	23
9 Zusammenfassende Bewertung	25
9.1 EXB-120 CHS	25
9.2 UniTree	25
9.3 Bewertungstabelle	26

10 Einsatzmöglichkeiten	27
--------------------------------	-----------

Verzeichnis der Abbildungen

Abb. 1	FTP/NFS Performance	7
Abb. 2	Migration bei mehreren Kopien	18
Abb. 3	Recall einer 10 MB Datei	19
Abb. 4	NFS Performance	19
Abb. 5	FTP Performance	20
Abb. 6	EXB 120 Bearbeitungszeiten	21
Abb. 7	Schreiben auf mehreren Laufwerken	22
Abb. 8	Lesen von mehreren Laufwerken	22

1.1 UniTree

UniTree ist ein System zur Datenverwaltung in einer verteilten Computerumgebung. UniTree kann große Datenmengen bei einem zentralen Fileserver verwalten, indem es diese in einem hierarchischen Filesystem abspeichert.

Als ein hierarchisches Filesystem verwaltet UniTree mehrere Speicherstufen, in der derzeitigen Implementierung zwei. Die erste Speicherstufe bilden Magnetplatten, die als Cache dienen. Alle unter UniTree erzeugten Dateien werden in diesem Cache angelegt. Nach einer vom Systemadministrator definierten Zeitspanne werden diese Dateien in die zweite Speicherstufe kopiert, welche aus Magnetbändern oder optischen Platten bestehen kann. Bei diesem Test wurde als zweite Stufe ein EXABYTE EXB-120 Cartridge Handling Subsystem verwendet. Nachdem eine Datei in die zweite Stufe kopiert wurde, kann sie im Cache gelöscht werden. Dies geschieht jedoch erst, wenn der Platz im Cache von anderen Dateien benötigt wird. Sobald eine Datei wieder angesprochen wird, wird sie vom Magnetband in den Cache kopiert, falls sie nicht dort noch vorhanden ist. Jeder Dateizugriff geschieht nur im Cache. Sollte eine Datei verändert werden, wird eine neue Kopie in der zweiten Speicherstufe angefertigt, die vorige Kopie ist dann nicht mehr zugreifbar. Wenn eine Datei gelöscht wird, besteht durch Einrichtung einer "Trash Can" die Möglichkeit, das sofortige Löschen zu verhindern. Die Datei wird stattdessen in das "Trash can" Verzeichnis aufgenommen, von wo sie nach einer vorher definierten Zeit (z. B. 24 Stunden) gelöscht wird. Während dieser Zeitspanne kann der Benutzer die Datei wiederherstellen, indem er sie in ein anderes Verzeichnis kopiert. Aus Sicherheitsgründen können in der zweiten Stufe mehrere Kopien auf verschiedenen Bändern angefertigt werden.

Der Benutzer kann auf das UniTree System über Network File System (NFS) oder File Transfer Protocol (FTP) zugreifen, es gibt keinen eigenen UniTree Client.

1.2 EXB-120 Cartridge Handling Subsystem

Das EXABYTE EXB-120 Cartridge Handling Subsystem (CHS) ist ein Robotersystem welches EXABYTE Bandkassetten verwaltet und diese automatisch in eingebauten EXABYTE Laufwerken montiert.

Das EXB-120 CHS kann bis zu 116 8mm Datenkassetten und 1 bis 4 Bandlaufwerke vom Type EXABYTE EXB-8500 Cartridge Tape Subsystem (CTS) enthalten. Mit diesen Laufwerken können bis zu 5 GByte an Daten auf jede Kassette geschrieben werden, so daß das EXB-120 CHS eine Gesamtkapazität von 580 GByte bietet.

2 UniTree Installation

2.1 Aufwand

Die Installation der UniTree Software war sehr einfach. Mit dem AIX Tool *smit* wurde die UniTree Version 1.1 von Disketten aus installiert. Die gesamte Installation (Version 1.0, EXABYTE Treiber, Upgrade auf Version 1.1, Performance Patch) dauerte etwa eine Stunde. Es traten keine Probleme auf. Anschließend wurden 20 EXABYTE Kassetten für UniTree initialisiert, dies dauerte etwa 20 Minuten. Mit der UniTree Konfiguration, für die etwa 30 Minuten benötigt wurden, war die Installation abgeschlossen. UniTree ließ sich problemlos starten und konnte sofort benutzt werden.

2.2 AIX Abhängigkeiten

UniTree wurde auf einem IBM RISC System/6000 Mod. 32H mit 16 MB Hauptspeicher und zwei Laufwerken mit je 304 MB installiert. Die Betriebssystemversion war AIX 3.2.*. UniTree ließ sich problemlos installieren und hatte keinerlei Probleme mit dieser Konfiguration.

Wenn UniTree gestartet ist, muß der UniTree NFS Server ebenfalls gestartet sein. Dieser benutzt die default NFS Port Nummer, so daß der AIX NFS Server nicht mehr laufen kann. Da mit einem NFS Server keine zwei verschiedenen logischen Filesysteme verwaltet werden können, und sowohl das standard AIX Filesystem als auch das UniTree Filesystem logische Filesysteme sind, kann kein Teil des AIX Filesystems über NFS exportiert werden.

Der UniTree FTP Daemon kann mit einer anderen Port Nummer als der AIX FTP Daemon gestartet werden, so daß über FTP auf beide logischen Filesysteme zugegriffen werden kann, wenn der entsprechende Port angegeben wird (*ftp hostname port*).

2.3 Benutzereinrichtung

Damit Benutzer mit FTP auf ihre Daten unter UniTree zugreifen können, müssen Sie vorher dem UniTree System bekannt gemacht werden. Gleichzeitig wird damit das "Trash Can" Verzeichnis eingerichtet. Wenn Benutzer nur über NFS auf UniTree zugreifen wollen, ist die Einrichtung einer UniTree Benutzernummer nicht unbedingt notwendig, es gibt dann allerdings auch kein "Trash Can" Verzeichnis zum verzögerten Löschen.

Benutzer können entweder mit dem AIX Tool *smit* oder durch Editieren der Datei *"/usr/lpp/UniTree/etc/adduser"* und anschließendem Aufruf der UniTree Prozedur *unsau* eingerichtet werden. Die Benutzerkennungen müssen vorher schon im AIX eingerichtet worden sein. Es müssen dann der login Name, die UID, die GID und das Home-Directory von AIX übernommen werden. Unter UniTree wird dann ein Home-Directory im UniTree Filesystem eingerichtet. Damit hat der Benutzer auf dem UniTree Server zwei gleichnamige Home-Directories, eins im AIX Filesystem und eins im UniTree Filesystem. Jeder UniTree Benutzer muß also auch eine AIX Benutzerkennung auf dem UniTree Server erhalten. Er erhält damit einen AIX login Namen und ein dazugehöriges AIX Password, mit denen er mit FTP auf seine UniTree Daten zugreifen kann. Unter

NFS kann er mit der im AIX eingetragenen numerischen UID auf seine Daten zugreifen, wenn das UniTree Filesystem oder ein Teilbaum davon für das entsprechende Zielsystem exportiert worden ist.

3 UniTree Benutzerschnittstelle

3.1 Allgemeines

Es gibt keine spezielle Client-Software für UniTree, sondern UniTree benutzt die bei fast allen Systemen verfügbaren Komponenten *Network File System (NFS)* und *File Transfer Protocol (FTP)*. Dies hat den Vorteil, daß ein Client UniTree benutzen kann, ohne vorher irgendwelche zusätzliche Software installieren zu müssen, hat allerdings auch den Nachteil, daß die Schwächen von NFS und FTP in Kauf genommen werden müssen. Als Zusatz zu den normalen UNIX Filesystemmöglichkeiten gibt es eine sogenannte "Trash Can". Damit kann das Löschen von Dateien verzögert erfolgen, so daß bei versehentlichem Löschen die Datei innerhalb einer gewissen Zeitspanne wiederhergestellt werden kann.

3.2 Lokaler Zugang

Ein lokaler Zugriff, also ein Zugriff von Benutzerprozessen, die auf dem UniTree Server laufen, zu UniTree-Daten ist ebenfalls nur über NFS oder FTP möglich. Dies bedeutet einen erhöhten Aufwand und damit verringerte Zugriffsgeschwindigkeiten auch für auf dem UniTree Server laufende Anwendungen.

3.3 FTP Zugang

Damit ein Benutzer auf UniTree Daten mittels FTP zugreifen kann, muß dieser Benutzer auf dem UniTree Server eine Benutzernummer und ein dazugehöriges Paßwort erhalten. Mit dieser Benutzernummer kann er sich dann von jedem System bei dem UniTree Server einwählen und mit FTP-Kommandos auf seine Dateien zugreifen bzw. neue Dateien erstellen. Um Zugriff auf die Datei zu erhalten, muß stets die gesamte Datei übertragen werden.

Der Benutzer kann alle FTP-Kommandos verwenden, um seine Dateien unter UniTree zu verwalten. Zusätzlich bietet der UniTree FTP Server einige spezielle Befehle, die die Besonderheiten von UniTree abdecken. So können die "Trash Can" Parameter und die Anzahl der Kopien über FTP vom Benutzer geändert werden. Desweiteren kann er die Sicherheitsparameter (owner, group, permissions) einer Datei ändern und mit den Befehlen *stage* und *wait* die mehrstufige Speicherung berücksichtigen. Mit dem *dir* Befehl erhält der Benutzer auch darüber Auskunft, auf welcher Speicherstufe (disk oder archive) sich seine Datei augenblicklich befindet.

3.4 NFS Zugang

Eine weitere Zugangsmöglichkeit bietet NFS. Das gesamte UniTree Filesystem oder Teile davon können vom UniTree Server exportiert werden und dann bei einem berechtigten NFS-Client gemountet werden. Das UniTree Filesystem steht dann den Benutzern dieses Clienten wie ein lokales Filesystem zur Verfügung.

Der UniTree NFS Server nimmt keinerlei Abbildung der numerischen UserIdentification (UID) und der GroupIdentification (GID) vor, so daß der Benutzer mit der Client-UID

und Client-GID auf seine Dateien zugreift. Sollten diese nicht den IDs entsprechen, die beim UniTree Server eingetragen sind (also hier die der Rechenzentrums Konvention), kann er nicht auf seine Dateien im UniTree zugreifen. Es besteht jedoch bei entsprechenden Export Parametern die Möglichkeit, daß der Systemverwalter des Clients seinen Benutzern in einem gemounteten Teilbaum des UniTree Filesystems eigene Directories einrichtet. Die Benutzer können dann mit diesem Filesystem wie mit einem lokalen Filesystem arbeiten. Sie können dann jedoch nicht über FTP beim UniTree Server auf diese Dateien zugreifen.

Da der standard NFS Client verwendet wird, können keine Erweiterungen genutzt werden, die auf die mehrstufige Speicherung eingehen. So kann der Benutzer über NFS nicht feststellen, in welcher Speicherstufe sich seine Datei im Augenblick befindet. Desweiteren kann es, je nachdem wie das UniTree Filesystem gemountet wurde, Schwierigkeiten beim Zugriff auf die auf Band ausgelagerten Dateien geben, da die Anfangszugriffszeit sehr lang sein kann und sogenannte "time outs" auftreten.

3.5 Vergleich FTP — NFS

Aus Benutzersicht hat NFS viele Vorteile. Der wichtigste ist, daß es für den Benutzer so aussieht, als ob alle Dateien lokal vorhanden wären und er deshalb in seiner vertrauten Umgebung (bekannte Kommandos) arbeiten kann. Er muß allerdings Abstriche an der Performance, der Sicherheit und, wenn seine UID und GID nicht der des UniTree Servers entsprechen, an der globalen Verfügbarkeit seiner Daten machen.

Beim Zugriff über FTP ist eine höhere Geschwindigkeit zu erreichen, als über NFS, da die Datei komplett übertragen wird, und deshalb größere Blöcke übertragen werden können (siehe Abb. 1).

Wenn nur über FTP auf die UniTree Dateien zugegriffen wird, ist die Sicherheit gegen unberechtigtem Zugriff größer, da nur derjenige, der die Kombination von UniTree Login Name und Paßwort kennt, auf die Dateien zugreifen kann, wohingegen beim Zugriff über NFS zumindest der Systemadministrator des Client-Systems auch auf diese Dateien zugreifen kann.

Sollten die UID und die GID auf dem Client System nicht der UID und der GID auf dem UniTree Server entsprechen, ist der Zugang über NFS zwar eventuell möglich, aber auf diese per NFS erzeugten Daten kann dann nicht über FTP zugegriffen werden.

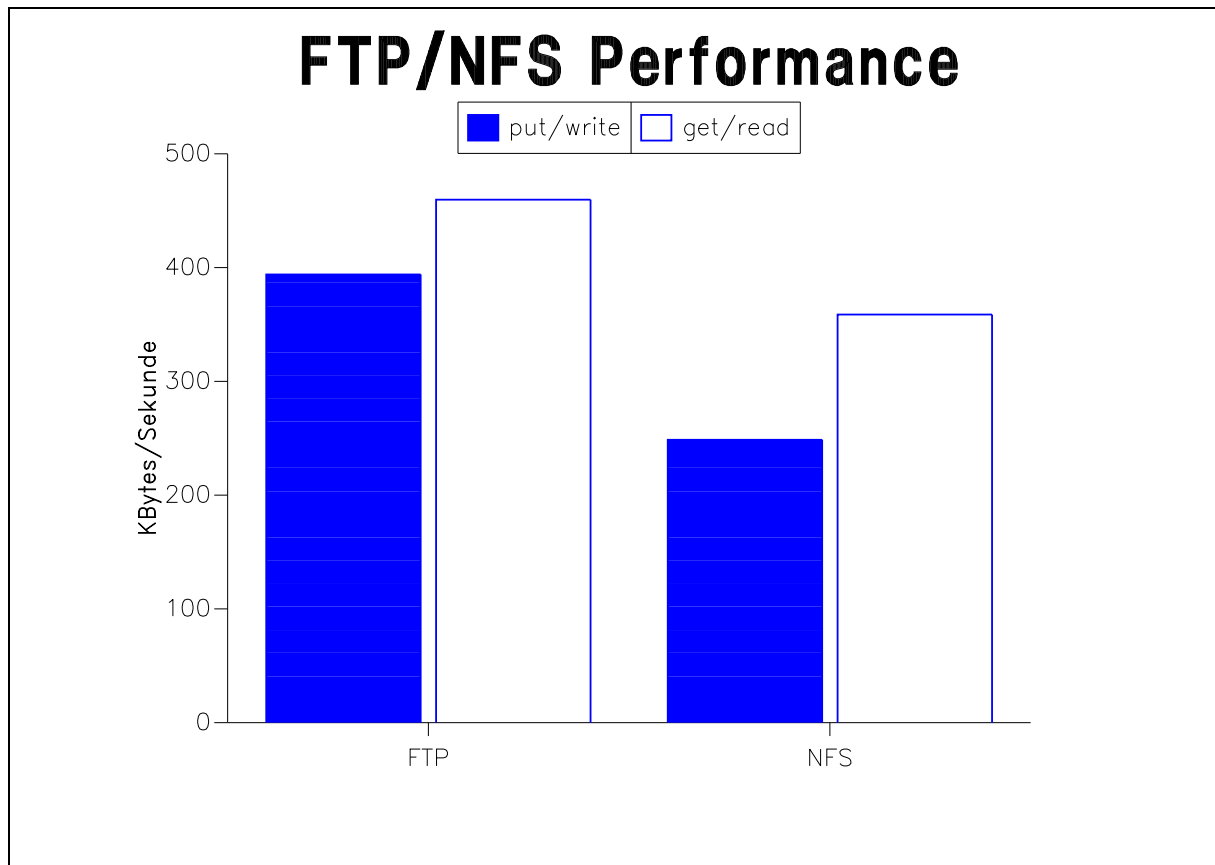


Abb. 1: FTP/NFS Performance

4 UniTree Systemadministration

4.1 Konfiguration

Konfigurationsänderungen werden mit dem AIX Tool *smit* vorgenommen. Nach der Änderung müssen die davon betroffenen Serverprozesse neu gestartet werden. Bei Änderungen, die die Bandkonfiguration betreffen, muß der Systemadministrator sicherstellen, daß zum Zeitpunkt der Änderung keine Bandaktivitäten stattfinden, da sonst Fehler auftreten, und der Benutzer eventuell I/O Fehler erhält bzw. wenn gerade Bänder beschrieben werden, der Rest des Bandes nicht mehr nutzbar ist. Dies wird leider nicht automatisch überprüft, und es gibt auch keine Möglichkeit, Bandaktivitäten zu verhindern bzw. zu beenden.

4.2 Benutzerverwaltung

UniTree Benutzer müssen, wenn sie über FTP auf ihre Daten zugreifen wollen, eingetragene AIX Benutzer sein. Die Eintragung erfolgt entweder über die AIX SMIT Schnittstelle oder durch Änderung spezieller Dateien. Zu Beachten ist, das als HOME Filesystem im UniTree der gleiche Name verwendet werden muß wie im AIX. Dies ist notwendig, damit per FTP zum HOME Filesystem verzweigt werden kann.

UniTree bietet keine Möglichkeit der Abbildung von numerischen Userids. Es muß deshalb eine homogene Benutzerkennung vorhanden sein.

Desweiteren ist es nicht möglich, den von einzelnen Benutzern belegten Platz zu kontingentieren. Dies ist aber unbedingt notwendig, da sonst einzelne Benutzer das System mit ihren Daten übermäßig belegen können und andere Benutzer dann keinen Platz mehr vorfinden.

4.3 Verwaltung der Kopien

UniTree bietet die Möglichkeit, mehrere Kopien einer Datei auf Band zu schreiben. Jede Kopie wird selbstverständlich auf ein anderes Band geschrieben. Normalerweise wird nur die erste Kopie beim Lesen verwendet. Durch (evtl. mehrmaligen) Aufruf des Administrator-Befehls *swapback* wird auf die jeweils nächste Kopie umgeschaltet, es erfolgt keine automatische Umschaltung. Solange kein Bandfehler auftritt könnten also die Bänder mit den weiteren Kopien aus dem Roboter entfernt werden und an sicherer Stelle aufbewahrt werden. Es gibt jedoch keine Möglichkeit, festzustellen, welche Bänder die Mehrfachkopien enthalten, so daß diese Bänder im Roboter verbleiben müssen. Dies ist aus Datenschutzgründen (Desaster Recovery) und Platzgründen (Verringerung der Roboterkapazität) untragbar.

4.4 Bandkompaktierung

Im Laufe der Zeit werden Dateien unter UniTree gelöscht bzw. geändert.. Dadurch wird der Belegungsgrad der Bänder mit aktiven Dateien immer geringer. Es gibt jedoch keine Möglichkeit, mehrere nur noch wenig belegte Bänder zu einem Band zusammenzufassen.

Ein Band kann nur dann wiederverwendet werden, wenn es vollständig leer ist. Es gibt auch keine Möglichkeit, festzustellen, welche Dateien noch auf einem Band sind, um diese eventuell manuell umzukopieren. Es ist daher zu erwarten, daß nach einiger Zeit nicht mehr die volle Kapazität des Roboters zur Verfügung steht. Im Extremfall kann die gesamte Kapazität durch 116 kleine Dateien (pro Band eine) erschöpft sein. Eine Möglichkeit zur Bandkompaktierung ist unbedingt notwendig.

4.5 Reinigung der Laufwerke

Nachdem ca. 200 GByte an Daten von den 4 EXABYTE Laufwerken verarbeitet wurden, sind diese mit einer Reinigungskassette gereinigt worden. Es wäre wünschenswert, wenn UniTree diese Reinigung automatisch vornehmen würde, sobald die in der Spezifikation der Laufwerke angegebenen Kriterien erfüllt sind. Für den Systemadministrator besteht keine Möglichkeit zu erkennen, wann diese Kriterien erfüllt sind. Er muß deshalb nach Gefühl die Reinigung durchführen oder abwarten, bis wegen aufgetretener Fehler eine Reinigung notwendig erscheint.

4.6 Sicherung der UniTree Datenbanken

Alle UniTree Datenbanken (Name-Server und Tape-Server) werden in zwei Kopien (mirror) gehalten, die auf verschiedenen Platten liegen sollten. Sollte eine der beiden Kopien ausfallen, muß diese Kopie mit Hilfe der anderen Kopie manuell wiederhergestellt werden. Sollten beide Kopien defekt sein, muß die Datenbank von Sicherungskopien wiederhergestellt werden. Sicherungskopien der Name-Server-Datenbank müssen manuell erstellt werden (mit *cp* oder *dd*). Es wird kein Datenbanklog mitgeführt, so daß alle Änderungen nach der Sicherung verloren sind. Sicherungskopien der Tape-Server-Datenbank werden mit Hilfe des UniTree Kommandos *backup* auf speziell gelabelte EXABYTE Kassetten (BCK000 bis BCK003) erstellt. Auch hier wird kein Datenbanklog mitgeführt, jedoch wird nach einer bestimmten Anzahl von Datenbankänderungen (nicht konfigurierbar) automatisch eine Zwischensicherung durchgeführt.

Diese Sicherungsmethoden sind völlig unzureichend. Notwendig wären

- automatische Umschaltung auf die zweite (mirror) Kopie,
- regelmäßige automatische inkrementelle Sicherung sowohl der Name-Server-Datenbank als auch der Tape-Server-Datenbank und
- Mitführen von Datenbankprotokollen, um mit der inkrementellen Sicherung und dem Datenbankprotokoll jederzeit den aktuellen Datenbankstand wiederherzustellen.

4.7 Beenden von UniTree

Es gibt keine Prozedur, die ein geregeltes Beenden von UniTree ermöglicht. Es besteht nur die Möglichkeit, alle UniTree Prozesse zu beenden (kill). Sollten zu diesem Zeitpunkt gerade Bänder beschrieben werden, tritt beim weiteren Beschreiben nach dem erneuten Start ein Schreibfehler auf und das Band wird als *voll* angesehen und kann dann nicht weiter beschrieben werden. Dadurch wird die Kapazität stark verringert (durchschnittlicher Verlust pro Band 2,5 GByte !). Sollten gerade Bänder gelesen werden, kann dies

nicht zu Ende geführt werden und der Benutzer bekommt einen I/O Fehler. Es besteht auch keine Möglichkeit, die Benutzer über die geplante Unterbrechung zu unterrichten.

5 UniTree Stabilität

Während der Testzeit (4 Wochen) lief die UniTree Software stabil und zuverlässig. Es wurden insgesamt 350 GByte an Daten erzeugt und zum großen Teil auch wieder gelesen und kopiert. Es sind keine Daten verloren gegangen. Außer den zu Testzwecken provozierten Fehlern traten keine Fehler auf. Nach den *Testfehlern* war jedoch öfters ein manueller Eingriff notwendig bzw. es traten Folgefehler auf. Insbesondere bei der Behandlung von Bandfehlern ist noch nicht die notwendige Robustheit vorhanden (siehe UniTree Error Recovery).

6 UniTree Error Recovery

6.1 Bandfehler

Um das Verhalten von UniTree bei einem Bandfehler zu testen, wurde ein aktives UniTree Band durch ein leeres Band ersetzt. Sobald eine Datei angesprochen wurde, deren primäre Migrationskopie auf diesem Band war, wurde das leere Band gemounted. UniTree merkte zwar, daß das Band nicht zu lesen war, es wurde jedoch keine automatische Fehlerkorrektur (Benutzung der zweiten Kopie dieser Datei, im Bandkatalog das Band als unbrauchbar vermerken) vorgenommen. Der Benutzerprozeß, der auf die Daten zugreifen wollte, ließ sich nicht mehr aus dem System entfernen.

Auf Dateien, die sich auf dem *defekten* Band befanden, konnte erst zugegriffen werden, nachdem durch den Administratorbefehl *swapback* auf die zweite Kopie umgeschaltet und UniTree neu gestartet wurde. Es gibt keine Möglichkeit, alle vom Bandfehler betroffenen Dateien festzustellen und dann auf die vorhandene Kopie umzuschalten. Dies ist jedoch eine unbedingt notwendige Funktion, um einen reibungslosen Betrieb auch bei Bandfehlern zu gewährleisten, insbesondere wegen der hohen Kapazität einer EXABYTE Kassette, auf der sich mehrere tausend Dateien befinden können.

Auch nachdem per Administratorbefehl *setblk tape -1* das Band für UniTree als nicht mehr verfügbar gekennzeichnet wurde, versuchte UniTree weiter auf dieses Band zuzugreifen, ebenso nachdem mit *setblk tape 0* das Band als leer gekennzeichnet wurde. Mit dem fehlerhaften Band war kein geordneter Betrieb mehr möglich, da immer wieder auf Dateien, die sich auf diesem Band befanden zugegriffen wurde.

Als nächstes wurde das gerade aktuelle Migrations-Band durch ein leeres Band ersetzt. UniTree bemerkte das fehlerhafte Band und stellte daraufhin die Migration ein, anstatt ein neues Migrations-Band zu benutzen.

Als weiterer Test wurde das gerade aktuelle Migrations-Band mit einem neuen (gleichen) Label versehen. Auch hier bemerkte UniTree, daß nicht zu dem aktuellen Block vorgespult werden konnte. In diesem Fall wurde mit der Migration auf ein neues Band begonnen.

Zusammenfassend kann gesagt werden, daß UniTree Fehler, die beim Schreiben eines Bandes vorkommen, erkennt und daraufhin mit einem neuen Band beginnt, daß aber bei Verlust oder Defekt eines kompletten Bandes, ein geordneter UniTree Betrieb nicht mehr möglich ist. Dazu müssen noch unbedingt die entsprechenden Werkzeuge (welche Dateien befinden sich auf einem Band und automatischer Wechsel zur nächsten Kopie) zur Verfügung gestellt werden.

6.2 Systemzusammenbruch

Um das Verhalten von UniTree nach einem Systemzusammenbruch zu testen, wurde die System Reset Taste gedrückt, während UniTree aktiv war (auch Bandaktivitäten). Nach dem System Start nahm UniTree den Betrieb wieder auf, es verblieb jedoch ein Band im Laufwerk, so daß ein manueller Eingriff notwendig wurde. Wenn zum Zeitpunkt des Zusammenbruchs gerade ein Band beschrieben wurde, wird nach dem Restart beim weiteren Beschreiben dieses Bandes ein Schreibfehler festgestellt und das Band wird all

voll angesehen (dieses Verhalten tritt auch beim Beenden von UniTree auf, wenn dann gerade ein Band beschrieben wird). Damit kann der Vorrat an leeren Bändern sehr schnell zu Ende gehen. Dies ist umso schlimmer, da keine Möglichkeit besteht, *fast* leere Bänder auf ein neues Band umzukopieren, um die freigewordenen Bänder wiederverwerten zu können.

7 UniTree Performance

Bei den Performance Untersuchungen sind zwei Aspekte wichtig,

- die interne UniTree Performance, das heißt, wie schnell kann UniTree die Daten auf Band migrieren und zurückholen
- die externe UniTree Performance, das heißt, wie schnell kann der Benutzer auf seine Daten zugreifen, wenn sie sich im UniTree Cache befinden.

7.1 interne Performance

Nachdem eine Datei im Cache angekommen ist, wird diese von UniTree ein oder mehrmals auf Band kopiert (migriert). Sobald alle Kopien erzeugt wurden, kann die Datei, wenn der Platz von anderen Dateien benötigt wird, aus dem Cache gelöscht werden. Es gibt drei Parameter die den Zeitpunkt der Migration bestimmen.

- Alter der Datei
Dieser Parameter gibt an, wie alt eine Datei sein muß, bevor sie migriert werden kann. Der Parameter darf nicht zu klein gewählt werden, da sonst die Möglichkeit besteht, daß bei kürzeren Netzausfällen über NFS erzeugte Dateien migriert werden, bevor sie vollständig im UniTree angekommen sind. Die Datei muß dann nochmals migriert werden, wenn auch der letzte Teil der Daten angekommen ist.
- Anzahl der Dateien im Cache
Dieser Parameter gibt an, wieviele zu migrierende Dateien im Cache sein müssen, damit eine Migration beginnt.
- Maximale Zeit zwischen der Migration
Dieser Parameter gibt an, nach welchem Zeitintervall, nachdem eine Datei angelegt oder geändert wurde, eine Migration auf jeden Fall stattfindet.

Diese Parameter müssen entsprechend der Cache-Größe und dem Benutzerverhalten gesetzt werden.

Sobald UniTree feststellt, daß eine Migration notwendig ist, werden alle Dateien, die auf Grund ihres Alters migriert werden können, migriert. Erst wenn diese Migration (eventuell mehrere Kopien) beendet ist, kann eine neue Migration starten. Vorher wird jedoch das Band entladen und muß neu geladen und vorgespult werden. Von UniTree wird nur ein Ausgabeband für jede Kopie vorgehalten, so daß keine parallele Migration stattfinden kann. Sobald eine Kopie einer zu migrierenden Datei fertig ist, wird mit der nächsten Kopie dieser Datei begonnen, so daß auch Kopien nicht parallel erstellt werden (Abbildung 2).

Dies hat zur Folge, daß UniTree maximal mit der Geschwindigkeit eines EXABYTE Laufwerkes Dateien migrieren kann. Es dürfen im Durchschnitt nicht mehr Daten erzeugt bzw. geändert werden als migriert werden können, da sonst der Cache unweigerlich überläuft, d. h. es dürfen nicht mehr als 20 GByte Daten pro Tag erzeugt oder geändert werden ($250 \text{ KB/sec} * 86400 \text{ sec}$). Spitzen müssen vom Cache abgefangen werden, der entsprechend groß sein sollte.

Beim Zurückladen in den Cache können alle verfügbaren Laufwerke genutzt werden, wenn sich die Daten auf verschiedenen Bändern befinden. Die Rückholgeschwindigkeit

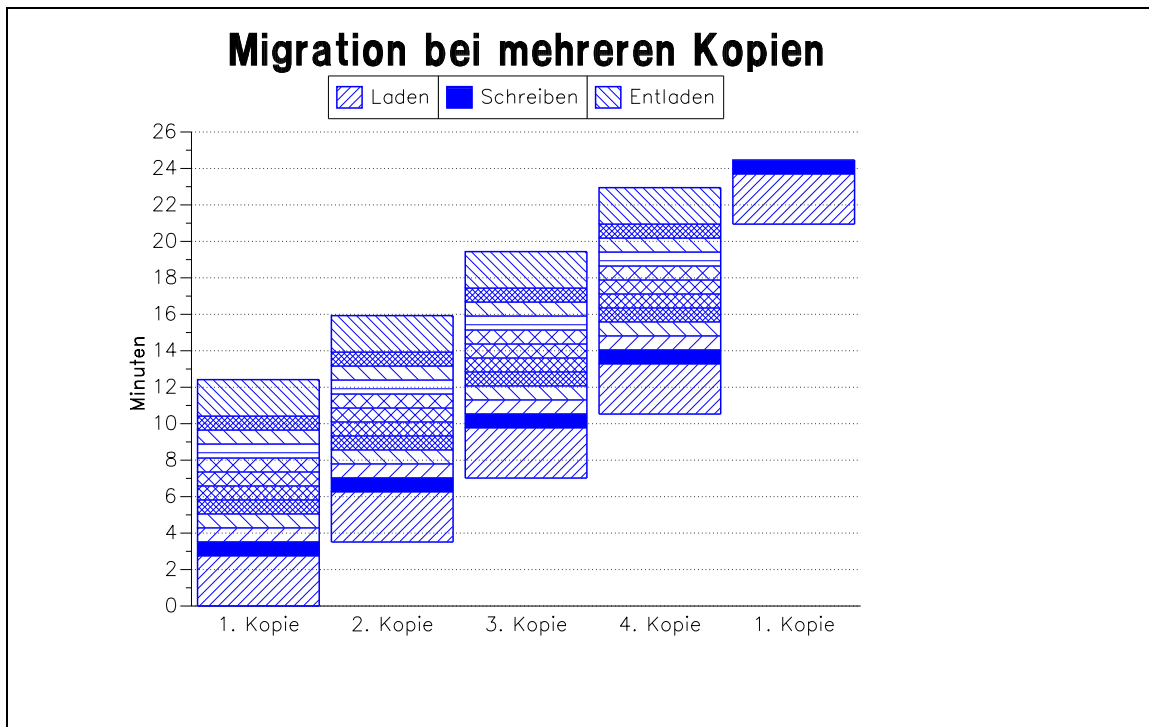


Abb. 2: Migration bei mehreren Kopien

hängt natürlich von der EXB-120 CHS Performance ab. Sie beträgt ca. 500 KB/sec pro Laufwerk, maximal jedoch 850 KB/sec bei Verwendung aller vier Laufwerke. Hinzu kommen die nicht unerheblichen Lade und Vorspulzeiten (siehe EXB-120 CHS Performance). Diese sind jedoch in Anbetracht der großen Kapazität einer Kassette hinnehmbar. Im Vergleich mit anderen Speichersystemen (HSM unter MVS und DMF unter UNICOS) sind die Recall-Zeiten nicht wesentlich schlechter (siehe Abbildung 3)

7.2 externe Performance

Die externe Performance wird bestimmt durch die möglichen Zugriffsmethoden FTP und NFS. Sobald die Datei im Cache verfügbar ist, kann der Benutzer auf die Daten zugreifen. Dabei sind die Zugriffszeiten, wie aus Abbildung 4 für NFS und Abbildung 5 für FTP ersichtlich, zwar etwas langsamer als die Zugriffszeiten auf AIX-interne Dateien (AIX-NFS und AIX-FTP), können sich jedoch im Vergleich mit den anderen Speichersystemen (UNICOS und MVS) durchaus sehen lassen.

Die lokalen Zugriffszeiten, also der Zugriff auf eine UniTree Datei durch auf dem UniTree Server laufende Benutzerprozesse, sind sehr langsam. Hier ist die RS/6000 320H deutlich überfordert, da ja sowohl der NFS-Server als auch der NFS-Client gleichzeitig benutzt werden. Jedoch ist die Performance für lokale Anwendungen im Normalfall von untergeordneter Bedeutung, da neben dem UniTree Server keine weiteren Anwendungen auf dem System laufen sollten. (Bei Verwendung des UniTree Systems als EXABYTE Server ist diese jedoch wichtig!)

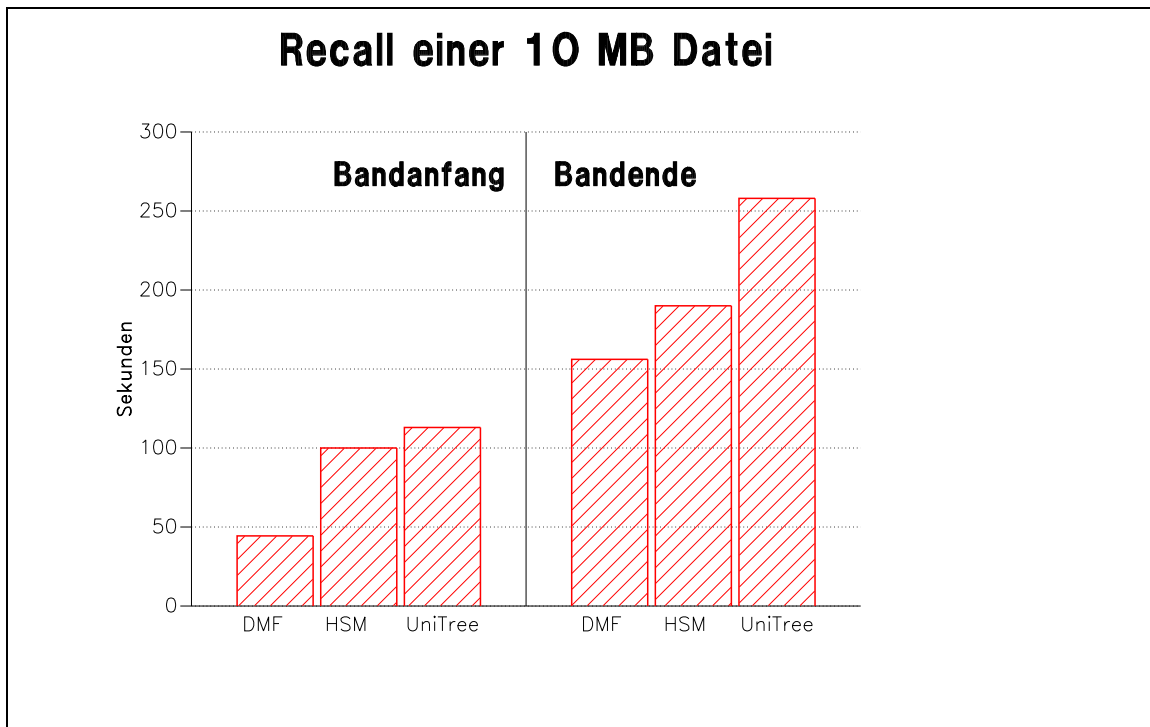


Abb. 3: Recall einer 10 MB Datei

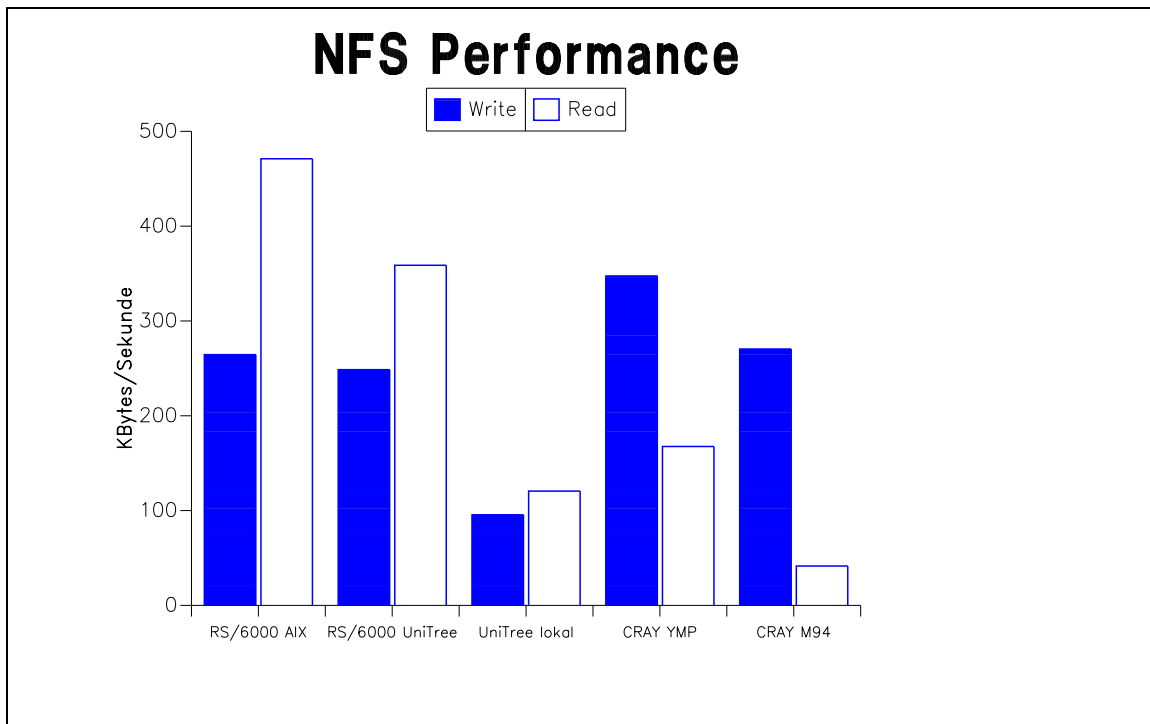


Abb. 4: NFS Performance

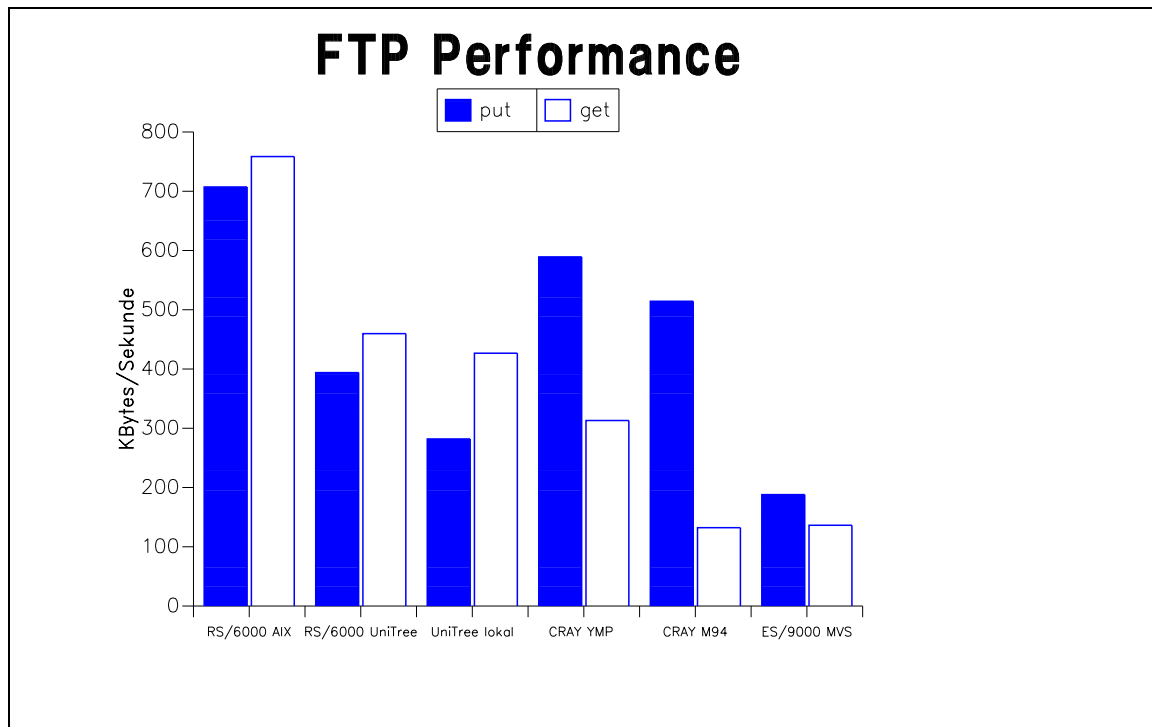


Abb. 5: FTP Performance

8 EXB-120 Cartridge Handling Subsystem

8.1 EXB-120 CHS Stabilität

Während der Testzeit traten keine Fehler auf. Alle Laufwerke und der Robotermechanismus arbeiteten einwandfrei. Es gab keine Probleme, ein auf einem Laufwerk beschriebenes Band auf einem anderen Laufwerk zu lesen. Nachdem 100 GByte an Daten auf Band geschrieben wurden und etwa ebensoviele Daten von Band gelesen wurden, sind alle 4 Laufwerke gereinigt worden (Reinigungskassette). Es waren jedoch noch keine Schreib- oder Lesefehler aufgetreten, die eine Reinigung notwendig gemacht hätten.

8.2 EXB-120 CHS Performance

Wie aus Abbildung 6 ersichtlich, ist die Schreib- und Lesegeschwindigkeit auf eine EXABYTE Kassette, die sich im EXB-120 CHS Kassettenroboter befindet von drei Faktoren abhängig. Da ist als erstes die Zeit, die vom Roboter benötigt wird, eine Kassette zum Laufwerk zu transportieren und dort einzulegen. Einschließlich der anschließenden Verifizierung des Labels dauert dieser Vorgang, wenn der Roboter frei ist, 1.5 Minuten. Anschließend muß dann zum Anfang der Datei vorgespult werden, was bis zu 2.5 Minuten dauern kann, wenn sich die Datei am Ende der Kassette befindet. Gelesen werden die Daten dann mit einer Geschwindigkeit von 400 bis 600 KB/sec, geschrieben mit 200 bis 300 KB/sec. Bei direkter Verwendung der EXABYTE-Laufwerke (ohne UniTree) ist eine höhere Geschwindigkeit zu erreichen (Schreiben bis 500 KB/sec und Lesen bis 700 KB/sec).

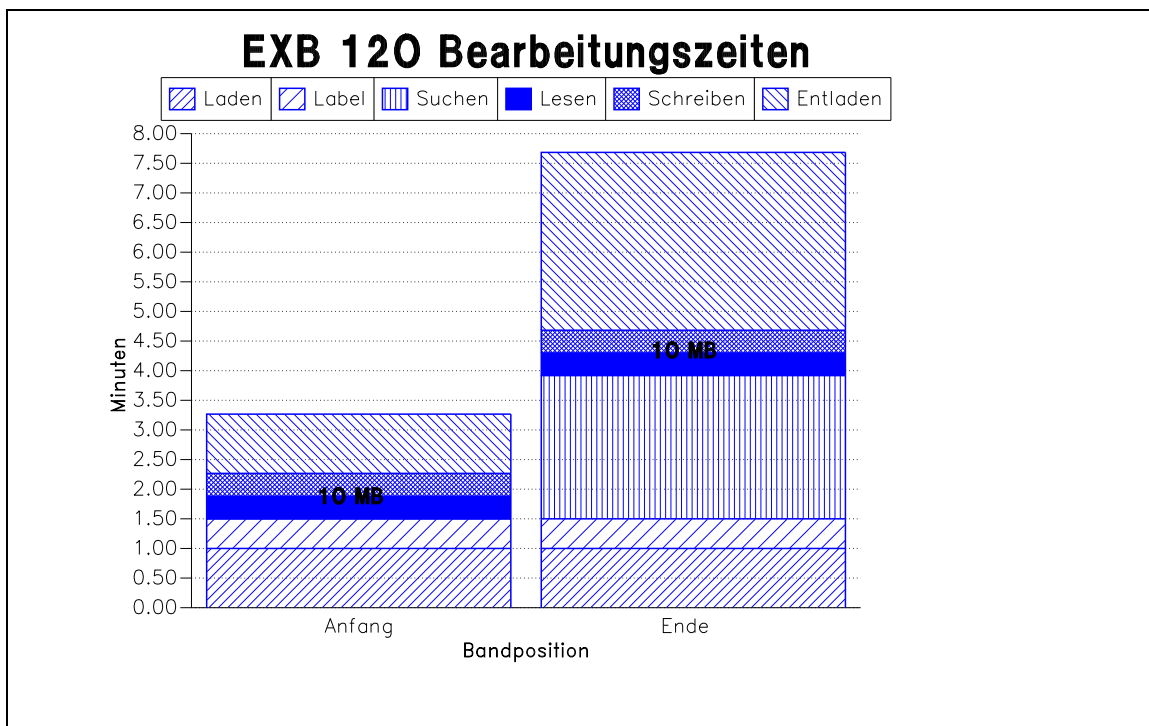


Abb. 6: EXB 120 Bearbeitungszeiten

Bei gleichzeitiger Benutzung mehrerer Laufwerke nimmt die erreichbare Geschwindigkeit pro Laufwerk ab (Abbildung 7 für Schreiben und Abbildung 8 für Lesen). Insgesamt können bei einer Ausnutzung aller vier Laufwerke etwa 700 KB/sec geschrieben und 850 KB/sec gelesen werden.

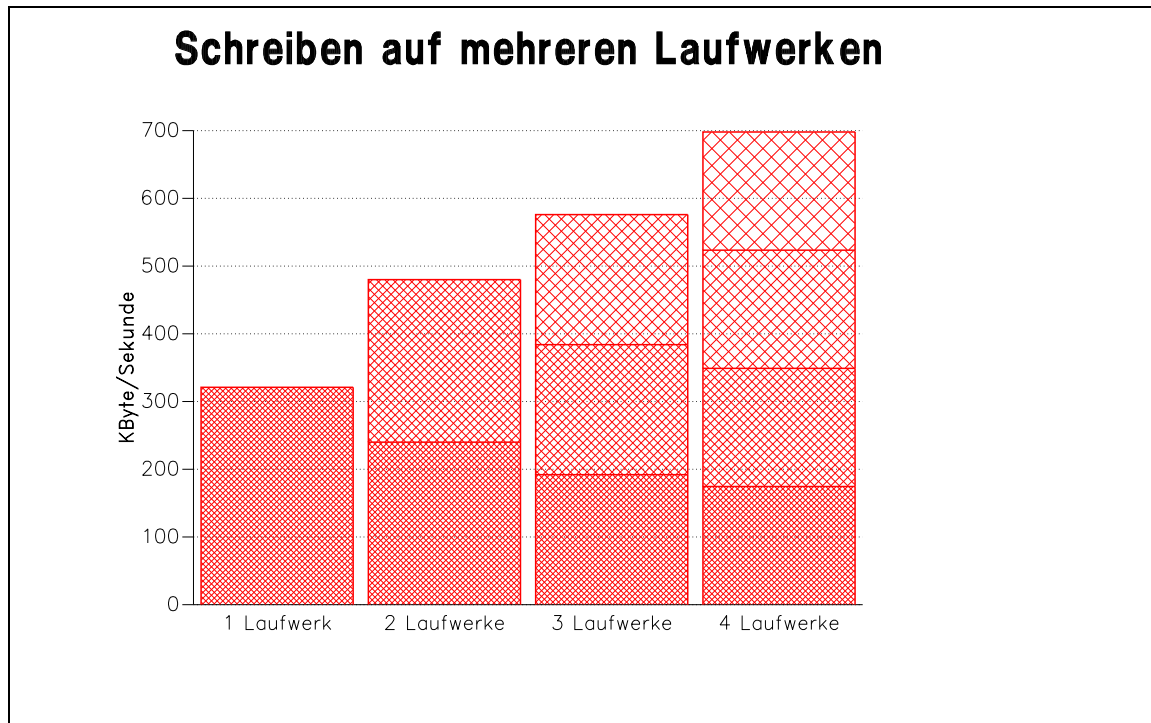


Abb. 7: Schreiben auf mehreren Laufwerken

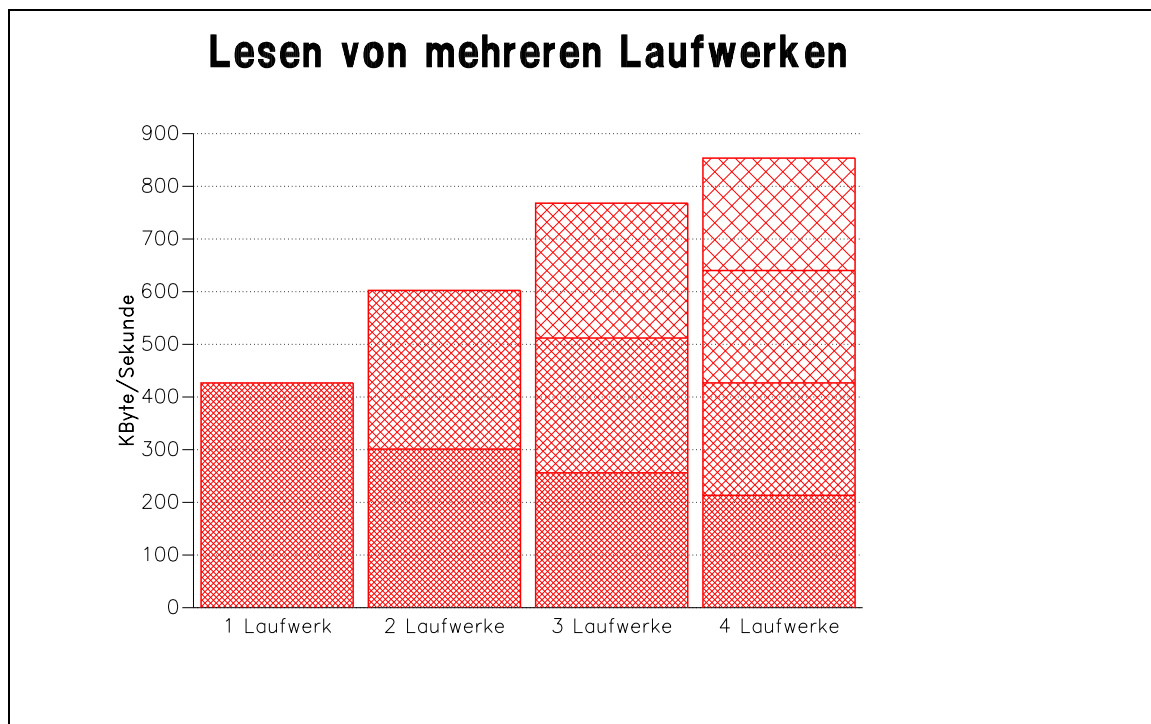


Abb. 8: Lesen von mehreren Laufwerken

8.3 EXB-120 CHS Roboterbefehle

Mit UniTree werden Kommandos ausgeliefert, die es ermöglichen das EXB-120 CHS zu steuern. So ist es möglich, Kassetten über den Entry/Exit-Port in das System einzugeben, diese dort abzustellen, zu den Laufwerken zu transportieren, wieder zurückzustellen und dann über den Entry/Exit-Port wieder auszugeben. Diese Funktionen könnten, nach Schaffung der entsprechenden Oberfläche dazu benutzt werden, Dateien aus dem UniTree System auf private EXABYTE Kassetten zu kopieren.

9 Zusammenfassende Bewertung

9.1 EXB-120 CHS

Während des gesamten Testes lief das EXB-120 CHS problemlos und ohne Fehler. Die zu erreichende Performance ist ausreichend und dem Preis der Laufwerke angemessen. Das EXB-120 CHS bietet auf wenig Raum zu günstigem Preis eine große Speicherkapazität. Wegen der Performance scheint dieses Gerät hauptsächlich für die Archivierung großer Datenmengen geeignet zu sein.

9.2 UniTree

Die dem UniTree zugrunde liegende Funktion, ein hierarchisches Speichersystem, wird erfüllt. Die Software läuft fehlerfrei und stabil, solange keine externen Fehler (Bandfehler) auftreten. Sollten jedoch Bandfehler auftreten ist der Systemadministrator mit der vorliegenden Version (UniTree 1.1) nicht in der Lage, angemessen auf diese Fehler zu reagieren. Sollte im Extremfall eine Kassette mit tausenden von Dateien unbrauchbar werden, ist das komplette UniTree System unbrauchbar. Hier müssen vor einem Produktionseinsatz noch unbedingt Verbesserungen stattfinden.

Die Zugangsmöglichkeiten für die Benutzer (NFS und FTP) haben den Vorteil, daß sie auf nahezu jedem System vorhanden sind, und daher auf UniTree von all diesen Systemen zugegriffen werden kann.

Der NFS Zugang hat jedoch Probleme mit der Sicherheit und der Abbildung heterogener Benutzerkennungen, so daß ein KFA-weiter Einsatz als zentraler Fileserver schon aus diesem Grund nicht möglich ist. Die fehlende Möglichkeit der Platzkontingentierung ist ein weiterer Hinderungsgrund. Mit NFS Zugang kann UniTree nur in einer kontrollierten Umgebung mit einheitlicher Benutzeradministration eingesetzt werden, sozusagen als virtuelles Filesystem für spezielle Rechner. In dieser Betriebsart hat UniTree den Vorteil gegenüber anderen Lösungen, daß Wartungszeiten synchronisiert werden können.

Bei ausschließlicher Benutzung des FTP Zugangs entfallen die Sicherheitsprobleme. Jedoch ist UniTree dann mehr ein Archivierungssystem, in dem große Datenmengen (Anzahl und Umfang der Dateien) archiviert werden können. Dem Vorteil, daß man dann von nahezu jedem System auf die Daten zugreifen kann, steht als Nachteil gegenüber, daß es keine Möglichkeiten gibt, zusätzlich zu dem Dateinamen noch weitere Informationen über den Inhalt der Datei im Archivierungssystem zu speichern.

9.3 Bewertungstabelle

EXB-120 CHS		
Stabilität		++ +
Performance		+
UniTree		
Installation		++ +
Benutzerschnittstelle		+
Stabilität		++ +
Robustheit		--
Sicherheit gegen unberechtigten Zugriff		-
Sicherheit gegen Datenverlust		++
Administration		---
	Sicherung der Datenbanken	--
	Zuordnung Band \leftrightarrow Datei	---
	Verwaltung der Bänder mit den Kopien	-
	Benutzerkontingentierung	---
	Geregeltes Herunterfahren von UniTree	--
	Bandkompaktierung	---
	Reinigung der EXABYTE Laufwerke	--
Performance		+

++ + sehr gut

--- unzureichend

10 Einsatzmöglichkeiten

Ein Einsatz in einer Produktionsumgebung ist erst dann möglich, wenn die fehlenden Administrator-Funktionen zu Verfügung stehen.

Wegen der Sicherheitsprobleme mit dem NFS Zugang kann UniTree auch dann nicht als zentraler Fileserver eingesetzt werden.

In der KFA könnte UniTree als virtuelles Filesystem für dedizierte Systeme wie z.B. Paragon oder das zentrale AIX eingesetzt werden. Der NFS Zugang ist in dieser Konfiguration sicher, da die gleichen zentralen Benutzerkennungen eingesetzt werden können. Die Zugriffsgeschwindigkeit dürfte schneller sein, als die augenblickliche CRAY NFS Geschwindigkeit. Abhängigkeiten von den CRAY Wartungszeiten entfallen.

Eine weitere Einsatzmöglichkeit wäre, UniTree als Archivierungssystem innerhalb der KFA einzusetzen. Als Zugangsmöglichkeit wird dann nur FTP angeboten. Bei diesem Einsatz steht UniTree in direkter Konkurrenz zur Archivierungskomponente des IBM Produktes ADSTM (ADSTAR Distributed Storage Manager). UniTree hat den Vorteil, daß es keine spezielle Client-Software benötigt. Es kann jedoch keine zusätzliche Information bezüglich des Dateiinhaltes mit abgespeichert werden, so daß der Benutzer Schwierigkeiten mit der Suche nach bestimmten Dateien hat. Außerdem bietet die Benutzeroberfläche der ADSTM Client-Software größeren Bedienungskomfort als das FTP. Dieser Einsatz kann auch in Verbindung mit dem Einsatz als virtuelles Filesystem für dedizierte Systeme angeboten werden.

Bei beiden Einsatzmöglichkeiten könnte UniTree zusätzlich für den automatisierten EXABYTE Kopierservice eingesetzt werden. Die Benutzer könnten ihre Daten über FTP zum UniTree Server bringen und dann dort lokal mit einem Batch Job diese Daten auf private EXABYTE Kassetten schreiben.